

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compen Ltd.

CM protein - protein search, using SW model

Run on: January 3, 2003, 13:01:31 ; Search time 20.3478 Seconds
(without alignments)
170.084 Million cell updates/sec

Title: US-09-801-784A-1

Perfect score: 173
Sequence: 1 VEKNITVTASVDPITDLOAGSALPSAVALTSPSPA 36

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

cheduled: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :
1: pfr73:*
2: pfr1:*
3: pfr2:*
4: pfr4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match length	ID	Description
1	155	89.6	170 1 YOEBC1	CFPI fimbrial prot
2	155	89.6	170 2 B56617	colonization facto
3	153	88.4	171 2 A41467	fimbrial protein c
4	138	79.8	170 2 S57935	ColA protein precu
5	113	65.3	191 2 AC0541	probable fimbrial
6	112	64.7	25 2 A64790	probable colonizat
7	72	41.6	20 2 A60100	fimbrial antigen C
8	60.5	35.0	74 2 S56208	hypothetical prote
9	58	33.5	1051 2 A39712	kinase-like protei
10	56.5	32.7	1580 1 EKECEX	gamma-glutamyltran
11	56.5	32.7	581 2 E91165	gamma-glutamyltran
12	56.5	32.7	581 2 E86011	gamma-glutamyltran
13	54	31.2	1186 2 T12737	tail protein - Met
14	53.5	30.9	417 2 D72372	hypothetical prote
15	53	30.6	360 2 S12850	protein TPX - Ther
16	53	30.6	474 2 S15921	protein TPX-VT3 -
17	53	30.6	740 2 A98160	hypothetical prote
18	53	30.6	740 2 H86005	hypothetical prote
19	52.5	30.3	243 2 E70846	hypothetical prote
20	52	30.1	498 2 G97279	protein containing
21	52	30.1	514 2 A31643	cell adhesion 80K
22	52	30.1	694 2 F97279	TPR-repeat-contain
23	52	30.1	740 1 B65136	YnfP protein - Esc
24	52	30.1	952 1 T46550	exonuclease ABC C
25	51.5	29.8	1708 2 AE1866	WD-40 repeat prote
26	51.5	29.8	415 2 A11611	competence-damage
27	51	29.5	833 2 S45042	genome polyprotein
28	50.5	29.2	194 2 G83915	hypothetical prote

30	50.5	29.2	580 2 AB0994	gamma-glutamyltran
31	50.5	29.2	3624 2 AD0835	large repetitive p
32	50.5	29.2	5175 2 T20992	hypothetical prote
33	50.5	29.2	5196 2 T43290	hemocentin precus
34	50	28.9	319 2 C96533	hypothetical prote
35	50	28.9	438 2 T37786	probable RNA-bind
36	50	28.9	848 2 T38089	HSP 70 family prot
37	49.5	28.6	163 2 AD0462	probable exported
38	49	28.3	50 2 F90761	hypothetical prote
39	49	28.3	138 2 S35233	CAP-1 protein - S
40	49	28.3	238 2 F91235	hypothetical prote
41	49	28.3	238 2 F86082	hypothetical prote
42	49	28.3	259 2 AE3185	dehydrogenase Atus
43	49	28.3	281 2 T51812	phosphoribosyl-AMP
44	49	28.3	299 2 S56031	pathogenesis-relat
45	49	28.3	449 2 F84241	hypothetical prote

ALIGNMENTS

RESULT 1
YOEBC1
CFPI fimbrial protein precursor - Escherichia coli
N:Alternate names: CFPI pilin; colonization factor antigen I (CFAI)
C:Species: Escherichia coli
C>Date: 14-Nov-1983 #sequence revision 30-Jun-1991 #text_change 15-Nov-1996
C:Accession: A30589; A03495; A43831
R:Kajalaenen, T.K.; Evans, D.G.; So, M.; Lee, C.H.
Index: Immun. 57, 1126-1130, 1989
A>Title: Molecular cloning and nucleotide sequence of the colonization factor antigen I
A:Reference number: A30589; MUID:89173309; PMID:2564374
A:Accession: A30589
A:Molecule type: DNA
A:Residues: 1-170 <KAR>
R:Klemm, P.
Eur. J. Biochem. 124, 339-348, 1982
A>Title: Primary structure of the CFPI fimbrial protein from human enterotoxigenic Esche
A:Reference number: A03495; MUID:82235736; PMID:6124420
A:Accession: A03495
A:Molecule type: DNA
A:Residues: 24-75, 'N', 77-96, 'A', 98-170 <KLE>
A:Experimental source: Strain H10407
R:Casellas, P.J.; Deal, C.D.; Reid, R.H.; Jarboe, D.L.; Nauss, J.L.; Carter, J.M.; Boedek
Index: Immun. 60, 2174-2181, 1992
A>Title: Analysis of Escherichia coli colonization factor antigen I linear B-cell epitop
A:Reference number: A43831; MUID:92267624; PMID:1375193
A:Accession: A43831
A:Molecule type: protein
A:Residues: 24-170 <CAS>
A:Experimental source: strain H10407
A:Note: sequence extracted from NCBI backbone (NCBI:104220)
C:Comment: The CFPI fimbriae are rather rigid, thread-like filaments of 0.5-1 micrometer
tical protein subunits.
C:Superfamily: CFPI fimbrial protein
C:Keywords: fimbria
F.1-23/Domain: signal sequence #status predicted <SIG>
F.24-170/Domain: CFPI fimbrial protein #status experimental <MAR>

Query Match 89.6% Score 155; DB 1; Length 170;
Best Local Similarity 88.9% Pred. No. 6.7e-14;
Matches 32; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 VEKNITVTASVDPITDLOAGSALPSAVALTSPSPA 36
|||
DB 24 VEKNITVTASVDPITDLOAGSALPSAVALTSPSPA 59
|||

RESULT 2
B56617
colonization factor antigen I precursor - Escherichia coli plasmid NTPI13
C:Species: Escherichia coli
C>Date: 05-Jan-1996 #sequence_revision 05-Jan-1996 #text_change 03-May-1996

C:Accession: B56617
 R:Jordi, B.J.; Willshaw, G.A.; van der Zeijst, B.A.; Gastra, W.
 DNA Seq. 2, 257-263, 1992
 A:Title: The complete nucleotide sequence of region 1 of the CFA/I fimbrial operon of hu
 A:Reference number: A56617; MUID:92329981; PMID:1352712
 A:Accession: B56617
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-170 <JOR>
 A:Experimental source: enterotoxigenic strain, CFA/I-ST plasmid NTP113
 A:Note: sequence inconsistent with nucleotide translation
 A:Note: sequence extracted from NCBI backbone (NCBIN:108960, NCBIP:108970)
 C:Genetics:
 A:Gene: cfab
 A:Genome: plasmid
 C:Superfamily: CFAI fimbrial protein

Query Match 89.6%; Score 155; DB 2; Length 170;
 Best Local Similarity 88.9%; Pred. No. 6.7e-14;
 Matches 32; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 VEKNITVTASVDPTIDLLQADGSLPSAVALTYSPA 36
 |||||
 Db 24 VEKNITVTASVDPTIDLLQADGSLPSAVALTYSPA 59
 |||||

RESULT 3
 A41467
 fimbrial protein csaA - Escherichia coli plasmid pRU405
 N:Alternate names: CS1 pilus major subunit
 C:Species: Escherichia coli
 C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 26-Aug-1999
 C:Accession: A41467; S19003
 R:Perez-Casal, J.; Swartley, J.S.; Scott, J.R.
 Infect. Immun. 58, 3594-3600, 1990
 A:Title: Gene encoding the major subunit of CS1 pili of human enterotoxigenic Escherich
 A:Reference number: A41467; MUID:91034170; PMID:1977705
 A:Accession: A41467
 A:Molecule type: DNA
 A:Residues: 1-171 <PER>
 A:Cross-references: GB:M37148; GB:M58550; NID:g145573; PIDN:AAA23596.1; PID:g145574
 R:Jordi, B.J.A.M.; van Vliet, A.H.M.; Willshaw, G.A.; van der Zeijst, B.A.M.; Gastra, W
 FEMS Microbiol. Lett. 80, 285-270, 1991
 A:Title: Analysis of the first two genes of the CS1 fimbrial operon in human enterotoxig
 A:Reference number: S19002
 A:Accession: S19003
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-171 <JOR>
 A:Cross-references: EMBL:X62879; NID:g41169; PIDN:CAA44673.1; PID:g41171
 C:Genetics:
 A:Gene: csaA
 A:Genome: plasmid
 C:Superfamily: CFAI fimbrial protein

Query Match 88.4%; Score 153; DB 2; Length 171;
 Best Local Similarity 83.3%; Pred. No. 1.3e-13;
 Matches 30; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 1 VEKNITVTASVDPTIDLLQADGSLPSAVALTYSPA 36
 |||||
 Db 24 VEKNITVTASVDPTIDLLQADGSLPSAVALTYSPA 59
 |||||

RESULT 4
 S57935
 CofA protein precursor - Escherichia coli
 C:Species: Escherichia coli
 C:Date: 13-Jan-1996 #sequence_revision 01-Mar-1996 #text_change 26-Aug-1999
 C:Accession: S57935
 R:Froehlich, B.J.; Karakashian, A.; Melsen, L.R.; Scott, J.R.
 Submitted to the EMBL Data Library, January 1995
 A:Description: The genes for CS2 pili of enterotoxigenic Escherichia coli and their inte

A:Reference number: S57934
 A:Accession: S57935
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-170 <PRO>
 A:Cross-references: EMBL:Z47800; NID:g597725; PIDN:CAA87761.1; PID:g597727
 C:Superfamily: CFAI fimbrial protein

Query Match 79.8%; Score 138; DB 2; Length 170;
 Best Local Similarity 79.4%; Pred. No. 1.5e-11;
 Matches 27; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 2 EKNIITVTASVDPTIDLLQADGSLPSAVALTYSP 35
 |||||
 Db 25 EKNIITVTASVDPTIDLLQADGSLPSAVALTYSP 58
 |||||

RESULT 5

AC0541

probable fimbrial chain tcfB [imported] - Salmonella enterica subsp. enterica serovar Ty
 C:Species: Salmonella enterica subsp. enterica serovar Typhi
 A:Note: this species has also been called Salmonella typhi
 C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 27-Nov-2001
 C:Accession: AC0541
 R:Farhnik, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
 th, T.; Conner, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
 S.; Moule, S.; O'Gaora, P.
 Nature 413, 848-852, 2001
 A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
 A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov

A:Reference number: AB0502; PMID:11677608

A:Accession: AC0541

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-191 <PAR>

A:Cross-references: GB:AL513382; PIDN:CAD08771.1; PID:g16501587; GSPDB:GN00176
 C:Genetics:
 A:Gene: tcfB
 C:Superfamily: CFAI fimbrial protein

Query Match 65.3%; Score 113; DB 2; Length 191;

Best Local Similarity 60.0%; Pred. No. 4.6e-08;

Matches 21; Conservative 9; Mismatches 5; Indels 0; Gaps 0;

QY 1 VEKNITVTASVDPTIDLLQADGSLPSAVALTYSP 35

 |||||
 Db 44 VOKDITVTANIDSTLELQADGSSLPSTNKLDPMP 78
 |||||

RESULT 6

A44790

probable colonization factor O166 - Escherichia coli (fragment)

C:Species: Escherichia coli

C:Date: 24-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 12-Apr-1995

C:Accession: A44790

R:Sommerfelt, H.; Grewal, H.M.; Svennerholm, A.M.; Gastra, W.; Flood, P.R.; Viboud, C.;
 Infect. Immun. 60, 3799-3806, 1992
 A:Title: Genetic relationship of putative colonization factor O166 to colonization facto

A:Reference number: A44790; MUID:92363580; PMID:1354200

A:Contents: E7476A

A:Accession: A44790

A:Status: preliminary

A:Molecule type: protein

A:Residues: 1-25 <SOM>

A:Note: sequence extracted from NCBI backbone (NCBIP:111010)

C:Superfamily: CFAI fimbrial protein

Query Match 64.7%; Score 112; DB 2; Length 25;

Best Local Similarity 92.0%; Pred. No. 6.1e-09;

Matches 23; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 VEKNITVTASVDPTIDLLQADGSL 25

 |||||
 |||||

Db 1 VEKNITVTSVDPITDILQNGSAL 25

RESULT 7

A:60100
C:Species: Escherichia coli (strain B8775) (fragment)
C:Date: 10-Nov-1992 #sequence_revision 10-Nov-1992 #text_change 18-Jun-1993
C:Accession: A60100
R:Wolfe, M.K.; Andrews, G.P.; Tall, B.D.; McConnell, M.M.; Levine, M.M.; Boedeker, E.C.
Infect. Immun. 57, 164-173, 1989
A:Title: Characterization of CS4 and CS6 antigenic components of PCF8775, a putative col
A:Reference number: A60100; MUID:89079281; PMID:2491834
A:Accession: A60100
A:Molecule type: protein
A:Residues: 1-20 <MOL>
C:Superfamily: CPAL fimbrial protein
Keywords: fimbria

Query Match 41.6%; Score 72; DB 2; Length 20;
Best Local Similarity 80.0%; Pred. No. 0.0015;
Matches 16; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 VEKNITVTSVDPITDILQ 20
| | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | |
Db 1 VEKNITVTSVDPITDILQ 20

RESULT 8

S56208
Hypothetical protein YFL047w - yeast (Saccharomyces cerevisiae)
C:Species: Saccharomyces cerevisiae
C:Date: 02-Sep-1995 #sequence_revision 19-Oct-1995 #text_change 19-Apr-2002
C:Accession: S56208
R:Wuzkang, Y.; Naitou, M.; Hagiwara, H.; Shibata, T.; Ozawa, M.; Sasamura, S.I.; Sasano
submitted to the EMBL Data Library, May 1995
A:Description: Analysis of the nucleotide sequence of chromosome VI from Saccharomyces ce
A:Reference number: S56186
A:Accession: S56208
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-714 <MUR>
A:Cross-references: EMBL:D50617; NID:9836685; PIDN:BA09194.1; PID:dl009832; PID:9836708
C:Genetics:
A:Gene: SGD:RGD2
A:Cross-references: SGD:S0001847
Map position: 6L

Query Match 35.0%; Score 60.5; DB 2; Length 714;
Best Local Similarity 35.7%; Pred. No. 3.6;
Matches 15; Conservative 7; Mismatches 13; Indels 7; Gaps 2;

QY 1 VEKNITVTSVDPITDILQ-----ADGSAFPSAVALT--KSP 35
| | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | |
Db 426 IDKMLTLENDIDPTADLLQLVKYKTSFKPQALVNNVNP 467

RESULT 9

A39712
Kinase-like protein klg precursor - chicken
C:Species: Gallus gallus (chicken)
C:Date: 08-Nov-1991 #sequence_revision 08-Nov-1991 #text_change 24-Sep-1999
C:Accession: A39712
R:Chou, Y.H.; Hayman, M.J.
Proc. Natl. Acad. Sci. U.S.A. 88, 4897-4901, 1991
A:Title: Characterization of a member of the immunoglobulin gene superfamily that possib
A:Reference number: A39712; MUID:91271300; PMID:1711213
A:Accession: A39712
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 11051 <CHO>
A:Cross-references: GB:M6337; NID:g212235; PIDN:AAA48933.1; PID:g212236
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homold

C:Keywords: ATP
F:775-1046/Domain: protein kinase homology <KIN>
F:783-791/Region: protein kinase ATP-binding motif

Query Match 33.5%; Score 58; DB 2; Length 1051;
Best Local Similarity 43.8%; Pred. No. 12;
Matches 14; Conservative 5; Mismatches 9; Indels 4; Gaps 1;

QY 3 KNITVTSVDPITDILQNGSALPSAVA 30
| | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | |
Db 499 KEVTVCSATGRKRTIQWTKIDSSLSFHSV 530

RESULT 10

EXCEX
gamma-glutamyltransferase (EC 2.3.2.2) precursor - Escherichia coli (strain K-12)
N:Alternate names: GGT protein; glutamyl transferase
N:Contains: gamma-glutamyltransferase large chain; gamma-glutamyltransferase small chain
C:Species: Escherichia coli
C:Date: 30-Jun-1991 #sequence_revision 30-Jun-1991 #text_change 01-Mar-2002
C:Accession: J00028; S47666; B65141; PC4349
R:Suzuki, H.; Kumagai, H.; Echigo, T.; Tochikura, T.
J. Bacteriol. 171, 5169-5172, 1989
A:Title: DNA sequence of the Escherichia coli K-12 gamma-glutamyltransferase gene, g
A:Reference number: J00028; MUID:89359163; PMID:2570061
A:Accession: J00028
A:Molecule type: DNA
A:Residues: 1-580 <SUZ>
A:Cross-references: EMBL:M28722; NID:9146132; PIDN:AAA23869.1; PID:9146133
R:Plunkett, G.
submitted to the EMBL Data Library, March 1994
A:Reference number: S47666
A:Accession: S47666
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-361 <PLU>
A:Cross-references: EMBL:U00039
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Ce
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617; PMID:9278503
A:Accession: B65141
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-580 <BLAT>
A:Cross-references: GB:AEO00421; GB:U00096; NID:91789854; PIDN:AACT6472.1; PID:91789856;
A:Experimental source: strain K-12, substrain MG1655
R:Hashimoto, M.; Suzuki, H.; Yamamoto, K.; Kumagai, H.
Biosci. Biotechnol. Biochem. 61, 34-39, 1997
A:Title: Analysis of low temperature inducible mechanism of gamma-glutamyltransferase
A:Reference number: PC4349; MUID:97179798; PMID:9028034
A:Accession: PC4349
A:Molecule type: DNA
A:Residues: 1-38 <HAS>
A:Experimental source: K-12 strain MG1655
C:Comment: This enzyme catalyzes the transfer of the gamma-glutamyl residue from gamma-g
utamyl compounds; however, its physiological role remains unclear. The active enzyme is

Query Match 32.7%; Score 56.5; DB 1; Length 580;
Best Local Similarity 40.0%; Pred. No. 10;
Matches 14; Conservative 9; Mismatches 7; Indels 5; Gaps 1;

QY 7 VTASVDPITDILQNGSALPSAVALTYS 36
| | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | |
Db 499 KEVTVCSATGRKRTIQWTKIDSSLSFHSV 530

Query Match 32.7%; Score 56.5; DB 1; Length 580;
Best Local Similarity 40.0%; Pred. No. 10;
Matches 14; Conservative 9; Mismatches 7; Indels 5; Gaps 1;

QY 7 VTASVDPITDILQNGSALPSAVALTYS 36
| | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | |
Db 499 KEVTVCSATGRKRTIQWTKIDSSLSFHSV 530

Query Match 32.7%; Score 56.5; DB 1; Length 580;
Best Local Similarity 40.0%; Pred. No. 10;
Matches 14; Conservative 9; Mismatches 7; Indels 5; Gaps 1;

QY 7 VTASVDPITDILQNGSALPSAVALTYS 36
| | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | |
Db 499 KEVTVCSATGRKRTIQWTKIDSSLSFHSV 530

Query Match 32.7%; Score 56.5; DB 1; Length 580;
Best Local Similarity 40.0%; Pred. No. 10;
Matches 14; Conservative 9; Mismatches 7; Indels 5; Gaps 1;

QY 7 VTASVDPITDILQNGSALPSAVALTYS 36
| | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | |
Db 499 KEVTVCSATGRKRTIQWTKIDSSLSFHSV 530

Query Match 32.7%; Score 56.5; DB 1; Length 580;
Best Local Similarity 40.0%; Pred. No. 10;
Matches 14; Conservative 9; Mismatches 7; Indels 5; Gaps 1;

QY 7 VTASVDPITDILQNGSALPSAVALTYS 36
| | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | |
Db 499 KEVTVCSATGRKRTIQWTKIDSSLSFHSV 530

Db 63 QSITITASSGTPMIDPTIALYNNNGSSY-SNLTIVGSP 99

RESULT 16

S15921 protein TPX-VT3 - Thermoproteus phage 1 (strain strain KRA1 10/12)

C:Species: Thermoproteus phage 1, TV1

A:Variety: strain KRA1 10/12

C:Date: 21-Nov-1993 #sequence_revision 26-May-1995 #text_change 20-Jun-2000

C:Accession: S15921; S06666

R:Neumann, H.; Zillig, W.

Nucleic Acids Res. 18, 2171, 1990

A:Title: Nucleotide sequence of the viral protein TPX of the TV1 variant VT3.

A:Reference number: S15921; MUID:90245666; PMID:2336394

A:Accession: S15921

A:Molecule type: DNA

A:Residues: 1-474 <NEW>

Cross-references: EMBL:X14717; NID:g62191; PIDN:CAA32838.1; PID:g584349

Experimental source: strain KRA1 10/12

Note: the authors translated the initiation codon GTG for residue 1 as Val

R:Neumann, H.

Submitted to the EMBL Data Library, March 1989

A:Reference number: S06686

A:Accession: S06686

A:Molecule type: DNA

A:Residues: 207-474 <NEW>

A:Cross-references: EMBL:X14717

A:Experimental source: strain KRA1 10/12

C:Genetics:

A:Start codon: GTG

C:Superfamily: human DNA-directed RNA polymerase II largest chain

Query Match 30.6%; Score 53; DB 2; Length 474;

Best Local Similarity 42.1%; Pred. No. 24;

Matches 16; Conservative 7; Mismatches 9; Indels 6; Gaps 2;

Db 3 KNITVTAS-----VDPTIDLQADGSAALPSAVALTYSPA 35

63 QSITITASSGTPMIDPTIALYNNNGSSY-SNLTIVGSP 99

RESULT 17

A98160

hypothetical protein ECs4249 [imported] - Escherichia coli (strain O157:H7, substrain R1)

C:Species: Escherichia coli

C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001

C:Accession: A98160

A:Variety: T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.

A:Variety: T.; Yasunaga, T.; Kunara, S.; Shibata, T.; Hattori, M.; Shinagawa, H.

DNA Res. 8, 11-22, 2001

A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and geno

A:Reference number: A98629; MUID:21156231; PMID:11258796

A:Accession: A98160

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-740 <RAY>

A:Cross-references: GB:BA000007; PIDN:BAH37672.1; PID:q13363723; GSPDB:GN00154

A:Experimental source: strain O157:H7, substrain R1MD 0509952

C:Genetics:

A:Gene: ECs4249

C:Superfamily: hypothetical protein ydcI

Query Match 30.6%; Score 53; DB 2; Length 740;

Best Local Similarity 37.5%; Pred. No. 40;

Matches 12; Conservative 4; Mismatches 16; Indels 0; Gaps 0;

Db 5 ITVTASVDPTIDLQADGSAALPSAVALTYSPA 36

84 IALEAGLEPLADLWSDPSHTPEVAAQYIDA 115

RESULT 18

H86005

hypothetical protein yhgF [imported] - Escherichia coli (strain O157:H7, substrain EDL93)

C:Species: Escherichia coli

C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001

C:Accession: H86005

R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew

Miller, L.; Grobeck, E.U.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,

Nature 409, 529-533, 2001

A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.

A:Reference number: A85480; MUID:21074935; PMID:11206551

A:Accession: H86005

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-740 <STO>

A:Cross-references: GB:AB005174; NID:g12518037; PIDN:AA659508.1; GSPDB:GN00145; UMGF:247

A:Experimental source: strain O157:H7, substrain EDL933

C:Genetics:

A:Gene: yhgF

C:Superfamily: hypothetical protein ydcI

Query Match 30.6%; Score 53; DB 2; Length 740;

Best Local Similarity 37.5%; Pred. No. 40;

Matches 12; Conservative 4; Mismatches 16; Indels 0; Gaps 0;

Db 5 ITVTASVDPTIDLQADGSAALPSAVALTYSPA 36

84 IALEAGLEPLADLWSDPSHTPEVAAQYIDA 115

RESULT 19

E70846

hypothetical protein RV3342 - Mycobacterium tuberculosis (strain H37Rv)

C:Species: Mycobacterium tuberculosis

C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999

C:Accession: E70846

R:Coile, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.

; Connor, R.; Davies, R.; Devlin, K.; Feldwell, T.; Gentles, S.; Hamlin, N.; Holloyd, S.

Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.

Nature 393, 537-544, 1998

A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrett, R.G.

A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome

A:Reference number: A70500; MUID:98295987; PMID:9634230

A:Accession: E70846

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-243 <COL>

A:Cross-references: GB:AL021841; GB:AL123456; NID:g1261517; PIDN:CAA17114.1; PID:e125117

A:Experimental source: strain H37Rv

C:Genetics:

A:Gene: RV3342

C:Superfamily: bioC homology

F:39-135/Domain: bioC homology <BIOC>

Query Match 30.3%; Score 52.5; DB 2; Length 243;

Best Local Similarity 45.2%; Pred. No. 13;

Matches 14; Conservative 5; Mismatches 11; Indels 1; Gaps 1;

Db 1 VKNITVTASVDPTIDLQADGSAALPSAVAL 31

57 VERGLDVA-VDPIDVLRALPQTVAL 86

RESULT 20

G97279

protein containing cell adhesion domain [imported] - Clostridium acetobutylicum

C:Species: Clostridium acetobutylicum

C:Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 14-Sep-2001

C:Accession: G97279

R:Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,

J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.

J. Bacteriol. 183, 4823-4838, 2001

A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium C

A:Reference number: A96900; MUID:21359125; PMID:21359325

A:Accession: G97279

EMBO J. 8, 371-377, 1989

A>Title: The contact site A glycoprotein of Dictyostelium discoideum carries a phospholipid anchor
A:Reference number: A56857; MUID:89251561; PMID:2721485
A:Contents: annotation; detection of glycosphingolipid:nitrosil anchor
A>Note: no phosphoserine was detected but phosphoinositol and ceramide were
A>Note: the N-linked carbohydrates are sulfated type I oligosaccharides
C:Comment: This protein is involved in the formation of intercellular contacts upon attachment.
C:Genetics:
A:Gene: csa

C:Keywords: blocked carboxyl end; cell adhesion; glycoprotein; lipoprotein; membrane protein
F:1-19/Domains: signal sequence #status predicted <SIG>
F:20-514/Product: cell adhesion 80K glycoprotein #status predicted <MAT>
F:132-139/Region: cell adhesion #status predicted
F:460-479/Region: 10-residue repeats
F:493-514/Domains: carboxyl-terminal propeptide #status predicted <cmp>
F:128_137_207_294_399/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:492/Modified site: GSI-anchor ethanolamine amidated carboxyl end (Ser) (in matrix form)

Query Match 30.1%; Score 52; DB 2; Length 514;
Best Local Similarity 26.5%; Pred. No. 36;
Matches 9; Conservative 7; Mismatches 18; Indels 0; Gaps 0;

QY 2 EKNIYTASVDPTIDLLQADGSALPSAVALTYP 35
|||::|
DB 251 ESSNTITAKASTGVDMIVLDNQGNOOPITFTYNP 284

RESULT 22

A44100

Cell adhesion molecule gp80 - slime mold (Dictyostelium discoideum)
C:Species: Dictyostelium discoideum
C:Date: 10-Jun-1993 #sequence_revision 18-Nov-1994 #text_change 18-Nov 1994
C:Accession: A44100
R:Desbarats, L.; Lam, T.Y.; Wong, L.M.; Siu, C.H.
J. Biol. Chem. 267, 19655-19664, 1992

A>Title: Identification of a unique cAMP-response element in the gene encoding the cell adhesion molecule gp80
A:Reference number: A44100; MUID:92406924; PMID:1326559
A:Contents: AX2
A:Accession: A44100
A>Status: preliminary
A:Molecule type: nucleic acid
A:Residues: 1-514 <DES>
A>Note: sequence inconsistent with the nucleotide translation
A>Note: sequence extracted from NCBI backbone (NCBI:113993)

Query Match 30.1%; Score 52; DB 2; Length 514;
Best Local Similarity 26.5%; Pred. No. 36;
Matches 9; Conservative 7; Mismatches 18; Indels 0; Gaps 0;

QY 2 EKNIYTASVDPTIDLLQADGSALPSAVALTYP 35
|||::|
DB 251 ESSNTITAKASTGVDMIVLDNQGNOOPITFTYNP 284

RESULT 23

F97279

TPR-repeat-containing protein, Cell-adhesion domain [imported]
C:Species: Clostridium acetobutylicum
C:Date: 14-Sep-2001 #sequence revision 14-Sep-2001 #text_change 14-Sep-2001
C:Accession: F97279
R:Nolling, J.; Bretton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee, J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001

A>Title: Genome Sequence and Comparative Analysis of the Solvent Producing Bacterium Clostridium acetobutylicum
A:Reference number: A96900; MUID:21359325; PMID:21359325
A:Accession: F97279
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-694 <Kur>
A:Cross-references: GB:AE001437; PIDN:AAK01025.1; PID:gl5026149; GSPUB:GN00168
A:Experimental source: Clostridium acetobutylicum ATCC2924
C:Genetics:
A:Gene: CAC3085

Query Match 30.1%; Score 52; DB 2; Length 694;
 Best Local Similarity 40.0%; Pred. No. 51;
 Matches 14; Conservative 7; Mismatches 10; Indels 4; Gaps 2;
 QY 1 VKNITVTASVDPPTIDLQADGSAALPSAVALTY 32
 DB 492 VDKNVTWTSS-DPSIATVAGSKITAVKPGTIVT 525

RESULT 24

B65136
 YnGF protein - Escherichia coli (strain K-12)
 C:Species: Escherichia coli
 C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 01-Mar-2002
 C:Accession: B65136
 R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co-
 lence, D.J.; Mau, B.; Shao, Y.
 Date: 277, 1453-1462, 1997
 Title: The complete genome sequence of Escherichia coli K-12.
 A:Reference number: A64720; MUID:97426617; PMID:9278503
 A:Accession: B65136
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Residues: 1-740 <BLAT>
 A:Cross-references: GB:AE000416; GB:U00096; NID:92367219; PID:AACT6432.1; PID:G1789811;
 A:Experimental source: strain K-12, substrain MG1655
 C:Genetics:
 A:Gene: ynfG
 C:Superfamily: hypothetical protein ydcI

Query Match 30.1%; Score 52; DB 1; Length 740;
 Best Local Similarity 37.9%; Pred. No. 55;
 Matches 11; Conservative 4; Mismatches 14; Indels 0; Gaps 0;

QY 5 ITVTASVDPPTIDLQADGSAALPSAVALTY 33
 DB 84 IATAGLEPLADLMDPSHTPEVAAQY 112

RESULT 25

T46550
 excinuclease ABC chain A [validated] - Thermus aquaticus
 N:Alternate names: nucleotide excision repair system protein uvrA
 N:Contains: excision endonuclease ABC (EC 3.1.-.-) chain A
 C:Species: Thermus aquaticus
 Date: 18-Feb-2000 #sequence_revision 18-Feb-2000 #text_change 19-Jan-2001
 Accession: T46550
 R:Yamamoto, N.; Kato, R.; Kuramitsu, S.
 Gene 171, 103-106, 1996
 A:Title: Cloning, sequencing and expression of the uvrA gene from an extremely thermophilic
 A:Reference number: Z23060; MUID:96257202; PMID:8675016
 A:Accession: T46550
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-952 <YAM>
 A:Cross-references: EMBL:P49111; PID:BAA08652.1
 A:Experimental source: strain HB8
 A:Note: the source is designated as Thermus thermophilus
 C:Comment: This protein is an ATPase and a DNA-binding protein that preferentially binds
 e, an ATP-dependent DNA repair enzyme that catalyzes the excision reaction of UV-damaged
 C:Genetics:
 A:Gene: uvrA
 C:Function:
 A:Description: plays an essential role in excision repair system [validated, MUID:962572
 A:Note: one of the most important DNA repair systems
 C:Superfamily: excinuclease ABC chain A; ATP-binding cassette homology
 C:Keywords: ATP; DNA binding; DNA repair; duplication; hydrolase; nucleotide binding; P-
 F;31-38/Region: nucleotide-binding motif A (P-loop)
 F:626-909/Domain: ATP-binding cassette homology <ABCE>
 F:643-650/Region: nucleotide-binding motif A (P-loop)

Query Match 30.1%; Score 52; DB 1; Length 952;

Best Local Similarity 35.6%; Pred. No. 74;
 Matches 16; Conservative 6; Mismatches 11; Indels 12; Gaps 2;
 QY 1 VKNITVTASVDPPTIDLQADGSAALPSAVALTYSP 35
 DB 890 IEHNDVVKTKADWDVLDLPGEGDRGGEIVAEOT--PEEVALTGP 932

Search completed: January 3, 2003, 13:02:31
 Job time : 22.3476 secs

